Some New Mixed Finite Elements in View of the Numerical Solution of Time Dependent Wave Propagation Problems

Eliane Bécache, Patrick Joly, Chrysoula Tsogka

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Some new mixed finite elements in view of the numerical solution of time dependent wave propagation problems

Eliane BECACHE, Patrick JOLY, Chrysoula TSOGKA

N° 3445

Juin 1998

THÈME 4

Rapport de recherche

ISSN 0249-6399

INRIA
ROCQUENCOURT

Some new mixed finite elements in view of the numerical solution of time dependent wave propagation problems

Eliane BECACHE*, Patrick JOLY†, Chrysoula TSOGKA‡

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet Ondes

Rapport de recherche n°3445 — Juin 1998 — 31 pages

Abstract: We present the construction and the analysis of a new family of quadrangular (2D) or cubic (3D) mixed finite elements, leading to an explicit scheme (mass lumping) for the approximation of the acoustic or elastic wave equations, including the case of an anisotropic medium. Non classical error estimates are given for this new element.

Key-words: mixed finite elements, mass lumping, anisotropic waves

(Résumé : tsvp)

* INRIA-rocquencourt. Eliane.Becache@inria.fr
† INRIA-rocquencourt. Patrick.Joly@inria.fr
‡ INRIA-rocquencourt. Chrysoula.Tsogka@inria.fr

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Nouveaux éléments finis mixtes pour la résolution numérique de problèmes de propagation transitoires

Résumé : Nous présentons la construction et l'analyse d'une nouvelle famille d'éléments finis mixtes quadrangulaires (2D) ou cubiques (3D), qui conduisent à des schémas explicites (condensation de masse) pour approcher les équations des ondes acoustiques ou élastiques dans des milieux anisotropes. Des estimations d'erreur non classiques ont été obtenues pour ce nouvel élément.

Mots-clé : éléments finis mixtes, condensation de masse, ondes anisotropes

Table of Contents

1 Introduction

This work falls in the more general framework of developing efficient numerical methods for approximating the equations of linear elastodynamics in complex media such as anisotropic heterogeneous media with cracks of arbitrary shapes. The main application of these studies is the non destructive testing. It consists in detecting the existence of cracks in a medium by studying the diffraction of a known incident wave by this medium. In the case we are interested in, the incident wave is a time pulse, which justifies a direct computation in the time domain.

The characteristics of this problem (large scale computations, complex geometries for the cracks, unbounded domains) and our ambition to construct an efficient method lead us to adopt the following criteria.

In order to facilitate the implementation and promote the speed of calculation we want to use regular meshes, squares in 2D and cubes in 3D. One can naturally think that this choice is restrictive and in contradiction with the fact that we want to take into account cracks of complex geometries.

In fact to deal with cracks we intend to use the fictitious domain method (cf. [21], [8]). Such a method allows us to work with uniform meshes independently of the geometry of the crack, the boundary condition being taken into account via the introduction of a Lagrange multiplier. Namely in the case of a crack in an elastic medium the boundary condition is a free surface condition, that means, the normal stress is zero on the crack : in order to consider this condition as an equality constant, we are led in a natural way, to the mixed velocity-stress formulation for elastodynamics. Then the Lagrange multiplier can be interpreted as the jump of the velocity through the crack.

In most examples the dimensions of the cracks are very small compared to the scale of the problem, which means that we have to model the elastic wave propagation in an unbounded domain. To do so we plan to use a new absorbing layer model, inspired by the Perfectly Matched Absorbing Layer (P.M.L) introduced by Berenger [5] for the 2D Maxwell problem. This model has astonishing properties : the reflection coefficient at the interface between the layer and the free medium is zero whatever are the frequency and the angle of incidence. The extension of this model to elastodynamics is natural when using the mixed velocity-stress formulation.

These considerations lead us to find an efficient approximation of the time domain mixed velocity-stress formulation. In particular for stability reasons (conservation of energy), we have decided to use a discretisation procedure in space based on a variational formulation of the velocity stress system, which is a first order hyperbolic system. At this stage, our main requirement is to define a spatial discretisation which allows the obtention of an explicit time discretisation scheme (mass-lumping).

Several mixed finite element methods are proposed in the literature especially for plane elasticity. We refer for example to PEER'S element introduced by D. N .Arnold, F. Brezzi and J. Douglas [1] and more recently to the work of R. Stenberg [19] and M. Morley [14]. One of the well known difficulties for mixed elements in elasticity consist in taking into account the symmetry of the stress tensor. The method used in ([1], [19], [14]) consist in working with a space of non necessarily symmetric tensors and imposing the symmetry in a weak way. Namely, the symmetry is enforced via the introduction of a Lagrange multiplier. Although these methods are very interesting for the plane elasticity problem, we did not retain them as they lead to an implicit scheme.

That is why we have constructed an original mixed finite element (inspired from the second Nédelec's family [16]) using spaces of symmetric tensors for the stress [3]. These spaces will fit our objectives.

The analysis of this new mixed finite element involves two main difficulties. The first one is due to the fact that the classical assumptions to get error estimates are not satisfied by this element, because, compared to classical approximations, we have enriched the approximate stress space but not the velocity one. The second difficulty in the analysis is linked to the symmetry of the stress tensor.

In this paper, we focus on the first difficulty. We present the new mixed finite element in the case of a model problem : the anisotropic wave equation. This equation can be seen as a simplified model for elastic waves in anisotropic media. We shall see that this simple case already poses new interesting theoretical questions we intend to solve here.

More precisely, in section 2, we will explain why the discretisation of this problem with the classical $RT_{[k]}$ mixed finite elements introduced by Raviart and Thomas in [17] does not lead to an explicit scheme and we will propose instead the use of a new mixed finite element. We first present the lowest order element in section 2.2, and then the extension to higher orders (2.3). Section 3 is concerned by the mixed approximation of the elliptic problem which is nothing but the stationary problem associated to the evolution problem of section 2. This analysis will be used to study an elliptic projection operator that will be useful for the analysis of the approximation of the evolution problem. We shall show in section 3.1 why the analysis of the new element does not fit the classical theory. That is why we develop in section 3.2 a new abstract theory leading to new error estimates. In section 3.3 we show that our new mixed finite elements enter this new framework and error estimates are given. Section 4 is devoted to relate the error estimates on the time domain solution to the error estimates obtained in the previous section on the elliptic problem. This essentially relies on energy estimates. Finally, section 5 presents the extension of the element to the 3D case.

In the present paper we have chosen to present the new element in the simplified case of the anisotropic wave equation in order to deal only with the first difficulty concerning mass lumping, some of the results we expose here have been announced in [4]. We will show in a next paper how we can generalize the family of the finite elements presented here in order to treat both difficulties (achieve mass lumping and take into account the symmetry of the stress tensor) at the same time.

2 Presentation of the new mixed finite elements

2.1 Position of the model problem : the anisotropic wave equation

Let Ω be a bounded domain of $\mathbb{R}^2$, $A(x)$ is a positive definite symmetric matrix satisfying :

$$A(x)\,\xi\cdot\xi \geq \alpha\,|\xi|^2 \qquad \alpha>0,\ \forall\xi\in\mathbb{R}^N,\ \text{p.p. } x\in\Omega. \tag{1}$$

We consider the scalar evolution problem

$$\left\{\begin{array}{l} \text{Find } u\ : [0,T]\rightarrow H_0^1(\Omega) \text{ such that } : \\ \\ \dfrac{\partial^2 u}{\partial t^2} - \text{div }(A^{-1}(x)\nabla u) = f, \qquad\qquad f\in C^0(0,T;L^2(\Omega)) \end{array}\right. \tag{2}$$

with the initial conditions

$$u(t=0)=u_0\in H^1\ ;\ \frac{\partial u}{\partial t}(t=0)=u_1\in L^2$$

Defining now

$$\left\{\begin{array}{l} p = A^{-1}(x)\nabla u \\ \\ v = \dfrac{\partial u}{\partial t} \end{array}\right.$$

and replacing it in (2) yields

$$(3) \qquad \begin{cases} \dfrac{\partial v}{\partial t} - \text{div } p = f \\[2mm] A\dfrac{\partial p}{\partial t} - \nabla v = 0 \end{cases}$$

with initial conditions

$$(4) \qquad p(0) = p_0 = A^{-1}(x)\nabla u_0 \ ; \ v(0) = v_0 = u_1$$

A mixed formulation associated to equations (3) is given by :

$$(5) \qquad \begin{cases} \text{Find } (p,v) : [0,T] \to X \times M \equiv H(\text{div }; \Omega) \times L^2(\Omega) \text{ such that :} \\[2mm] \dfrac{d}{dt}a(p,q) + b(v,q) \quad = 0 \qquad \forall q \in X \\[2mm] \dfrac{d}{dt}(v,w) - b(w,p) \quad = (f,w) \quad \forall w \in M, \end{cases}$$

where

$$(6) \qquad \begin{cases} a(p,q) = \displaystyle\int A(x)\ p \cdot q\ dx \quad \forall (p,q) \in X \times X \\[2mm] b(w,q) = \displaystyle\int w \ \ \text{div } q\ dx \quad \forall (w,q) \in M \times X \\[2mm] (f,w) = \displaystyle\int f\ w\ dx \qquad\qquad \forall w \in M. \end{cases}$$

The bilinear form $a(\cdot,\cdot)$ (resp. $b(\cdot,\cdot)$) is continuous on $H \times H (H = (L^2(\Omega))^2)$ (resp. on $X \times M$). The bilinear form $a(\cdot,\cdot)$ (resp. $b(\cdot,\cdot)$) thus defines a linear continuous operator $\mathcal{A} : H \to H'$ by $\langle \mathcal{A}p, q\rangle_{H' \times H} = a(p,q)$ (resp. $B : M \to X'$ by $\langle Bw, q\rangle_{X' \times M} = b(w,q)$. They satisfy the following properties (see for instance [7]) :

$$(7) \qquad \left| \begin{array}{ll} (i) & \text{The continuous inf-sup condition} \\[2mm] & \exists c > 0 \,/\, \forall\ w \in M,\ \exists\ q \in X \,/\, b(w,q) \geq c\, \|w\|_M \, \|q\|_X \\[2mm] (ii) & \text{The coercivity of the form } a(\cdot,\cdot) \text{ on } V (= \text{Ker} B) \\[2mm] & \exists \alpha > 0 \,/\, \forall\ p \in V = \{q \in X \,/\, b(w,q) = 0,\ \forall\ w \in M\}, \quad a(p,p) \geq \alpha\, \|p\|_X^2 . \end{array} \right.$$

In the following, we only consider the semi-discretisation in space of this problem, keeping in mind our main motivation which is the possibility to do mass lumping.

2.2 Presentation of the $Q_1^{div} - Q_0$ finite element in the lowest order

We suppose now that Ω is a union of rectangles in such a way that we can consider a regular mesh $(\mathcal{T}_h)$ with squares elements (K) of edge $h > 0$. We introduce the following approximation spaces :

$$(8) \qquad \begin{cases} X_h = \left\{ q_h \in X \,/\, \forall K \in \mathcal{T}_h,\ q_h|_K \in \hat{X} \right\} \\[2mm] M_h = \left\{ w_h \in M \,/\ \ \forall K \in \mathcal{T}_h,\ w_h|_K \in \hat{M} \right\} \end{cases}$$

where $\hat{X}$ (resp. $\hat{M}$) denotes a finite dimensional space of vector (resp. scalar) functions. The discrete problem associated to (5), (4) is

$$(9) \qquad \begin{cases} \text{Find } (p_h, v_h) : [0,T] \to X_h \times M_h \text{ such that :} \\[2mm] \dfrac{d}{dt}a(p_h,q_h) + b(v_h,q_h) \quad = 0 \qquad \forall q_h \in X_h \\[2mm] \dfrac{d}{dt}(v_h,w_h) - b(w_h,p_h) \quad = (f,w_h) \quad \forall w_h \in M_h, \end{cases}$$

with initial conditions

$$p_h(0) = p_{0,h} \ ; \ v_h(0) = v_{1,h}$$

The usual choice consists in taking for $\hat{X}$ the lowest order Raviart Thomas element :

$$\hat{X} = RT_{[0]} = P_{10} \times P_{01}$$

and for $\hat{M}$ piecewise constants :

$$\hat{M} = Q_0 \tag{10}$$

We remind here that P_k is the space of polynomials of degree $\leq k$ and $P_{k_1 k_2}$ is defined by :

$$P_{k_1 k_2} = \left\{ p(x_1, x_2) \ | \ p(x_1, x_2) = \sum_{i \leq k_1, j \leq k_2} a_{ij} x_1^i x_2^j \right\}$$

Let us explain now why this choice does not lead to an explicit scheme when one considers the evolution problem of anisotropic waves corresponding to (9). We introduce here $B_{N_1} = \{\tau_i\}_{i=1}^{N_1}$, $B_{N_2} = \{\phi_i\}_{i=1}^{N_2}$ the bases of X_h and M_h respectively ($N_1 = \dim X_h$ and $N_2 = \dim M_h$) and $[P] = (P_1, ..., P_{N_1})$, $[U] = (U_1, ..., U_{N_2})$ the coordinates of p_h, u_h in the bases B_{N_1}, B_{N_2}. In these bases, problem (9) can be written in the following form :

$$\begin{cases} \text{Find } (P, U) \in L^2(0, T; (\mathbb{R}^{N_1})) \times L^2(0, T; (\mathbb{R}^{N_2})) \text{ such that :} \\ M_p \dfrac{dP}{dt} + C^T U = 0 \\ M_u \dfrac{dU}{dt} - CP = F \\ + \text{ initial conditions} \end{cases}$$

with :

$$\begin{array}{lll} (i) & (M_p)_{i,j} = (A\tau_i, \tau_j)_{(L^2(\Omega))^2}, & 1 \leq i, j \leq N_1 \\ (ii) & (M_u)_{i,j} = (\phi_i, \phi_j)_{L^2(\Omega)}, & 1 \leq i, j \leq N_2 \\ (iii) & (C)_{i,j} = (\phi_i, \text{div}\tau_j)_{L^2(\Omega)}, & 1 \leq i \leq N_2, \ 1 \leq j \leq N_1 \\ (iv) & (F)_j = (f, \phi_j)_{L^2(\Omega)} & 1 \leq j \leq N_2 \end{array} \tag{11}$$

C^T denotes the transpose of C. If we use a centered finite difference approximation for the time discretisation, the solution at each time step is obtained by inverting the mass matrices M_u and M_p. Although they are symmetric and sparse, this inversion can become costly (for large systems) and we prefer to avoid it. In that order, we want to reduce them to diagonal (or block diagonal) matrices by using a mass lumping technique (see [13], [20]). This consists in using adequate quadrature formulas to approximate the integrals in (11 (i), (ii)). One can remark that M_h being chosen discontinuous, the matrix M_u does not really need to be mass lumped (it is already diagonal here) so we focus our attention on the mass lumping of M_p.

Let now (τ_i) be the $RT_{[0]}$ basis functions (see figure 1) and consider the integral

$$a(\tau_a, \tau_b) = \int_\Omega A\tau_a \cdot \tau_b dx = \sum_k \int_{K \in \mathcal{T}_h} A\tau_a \cdot \tau_b dx \tag{12}$$

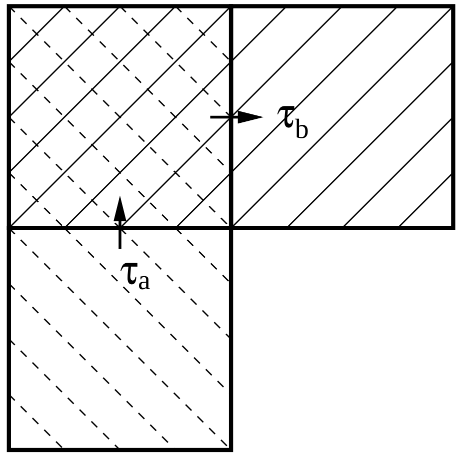

Figure 1: Two bases functions of $RT_{[0]}$

If we use the following quadrature formula :

$$\int_K f(x)dx = mes(K) \sum_{i=1}^{N} \omega_i f(M_i)$$

with $(M_i)_{i=1,...,N}$ the quadrature points and $(\omega_i)_{i=1,...,N}$ the associated weights to approximate (12), we obtain :

$$\int_K A\tau_a \cdot \tau_b dx = mes(K) \sum_i \omega_i A\tau_a(M_i) \cdot \tau_b(M_i)$$

This would lead to a diagonal matrix if

(13) $$A\tau_a(M_i) \cdot \tau_b(M_i) = 0 \quad \forall\ a \neq b$$

Consider the element K with the local enumeration (see figure 2) and $(\tilde{\tau}_i)_{i=1,...,4}$ the local $RT_{[0]}$ basis functions (associated to the wedges $(A_i)_{i=1,...,4}$ of the element K).

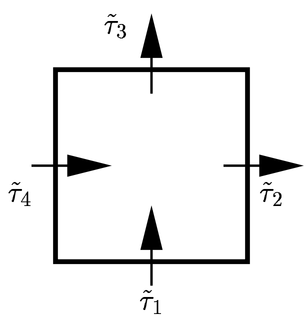

Figure 2: Local basis functions in $RT_{[0]}$

For the isotropic wave equation ($A = Id$) formula (13) reduces to

(14) $$\tilde{\tau}_a(M_i) \cdot \tilde{\tau}_b(M_i) = 0 \quad \forall\ a \neq b,\ a, b = 1, ..., 4$$

For two orthogonal edges, the associated basis functions are already orthogonal $((\tilde{\tau}_{2j+1}, \tilde{\tau}_j) = 0$ where the indexes are defined modulo 4).

Let us now consider the lowest order Gauss-Lobatto quadrature formula, using the summits of the element K as quadrature points,

$$\int_K f(x)dx = \frac{mes(K)}{4}\sum_{i=1}^{4} f(M_i) \tag{15}$$

it is easy to check that (14) is satisfied (since $\tilde{\tau}_j \equiv 0$ on A_{j+2}). On the contrary for the anisotropic wave equation, this does not work any more, since for two orthogonal edges, functions $A\tau_a$ and τ_b are not orthogonal any more. In fact, in this case, there exists no quadrature formula that satisfies (13) with the $RT_{[0]}$ basis functions.

The alternative solution that we propose consists in changing the approximation space $\hat{X}$. Let

$$\hat{X} = Q_1 \times Q_1 \tag{16}$$

this element was initially introduced by Nédelec in [16].We will call this choice (16) combined with (10) the $Q_1^{div} - Q_0$ element.

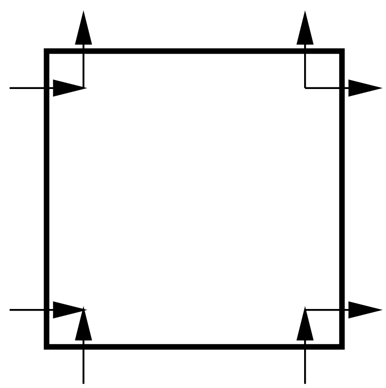

Figure 3: The $Q_1 \times Q_1$ element

The quadrature formula (15) satisfies now (13) with the new basis functions: the key point is that the quadrature points coincide with the degrees of freedom M_j and the new basis functions satisfy :

$$\tau_i(M_j) = \delta_{ij}$$

For more details on quadrature formulae and mass lumping techniques we refer on the work of G. Cohen, P. Joly and N. Tordjman for the acoustic wave equation ([20], [12]), to P. Monk, G. Cohen [13] and A. Elmkies [11] for Maxwell's equations.

Remark 1 *The mass matrix $a(p_h, q_h)$ is block diagonal, and after an inversion of local (4×4) matrices we obtain the explicit scheme.*

Before the analysis, let us show how this element can easily be extended to higher orders.

2.3 Extension to higher orders and mass lumping

The natural generalization of the lowest order element, presented in the previous section, consists in taking :

$$\left\{ \begin{array}{l} X_h = \{q_h \in X \;/\; \forall K \in \mathcal{T}_h,\; q_h|_K \in Q_{k+1} \times Q_{k+1}\} \\ \\ M_h = \{w_h \in M \;/\; \forall K \in \mathcal{T}_h,\; w_h|_K \in Q_k\} \end{array} \right. \tag{17}$$

and we will call it the $Q_{k+1}^{div} - Q_k$ element. This choice still satisfies our requirement with respect to mass lumping. We focus here on the mixed finite elements corresponding to $k = 1, 2$, that we use in practice. Presentation of the degrees of freedom in the reference element ($[0,1] \times [0,1]$) :

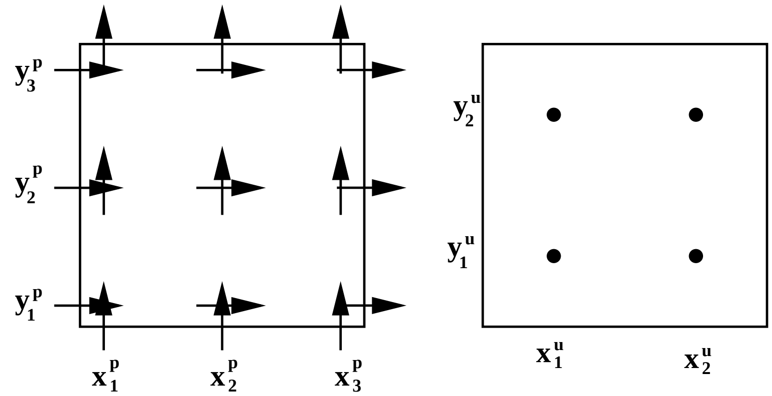

Figure 4: Degrees of freedom in the $Q_2^{div} - Q_1$ element

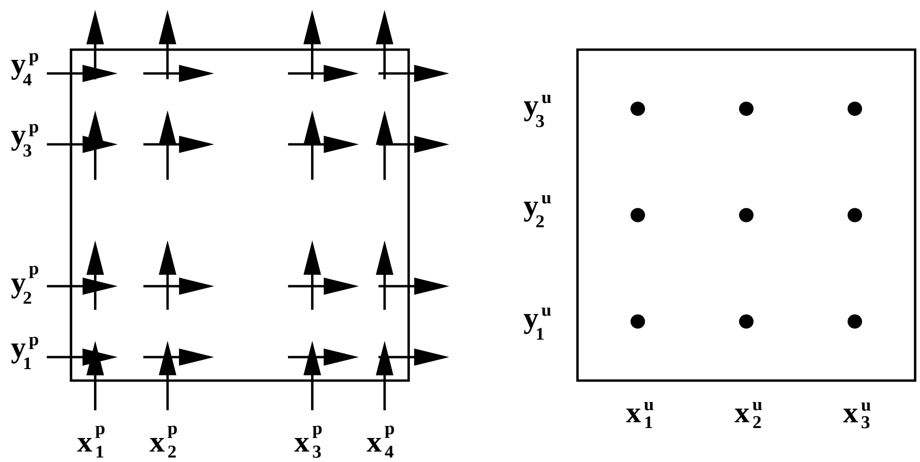

Figure 5: Degrees of freedom in the $Q_3^{div} - Q_2$ element

with

$$\text{for } k=1 \quad x_1^p = y_1^p = 0, \quad x_2^p = y_2^p = \frac{1}{2}, \quad x_3^p = y_3^p = 1$$

$$\text{for } k=2 \qquad x_1^p = y_1^p = 0, \qquad x_2^p = y_2^p = \frac{5-\sqrt{5}}{10}$$

$$x_3^p = y_3^p = \frac{5+\sqrt{5}}{10}, \quad x_4^p = y_4^p = 1$$

and

$$\text{for } k=1 \quad x_1^u = y_1^u = \frac{3-\sqrt{3}}{6}, \quad x_2^u = y_2^u = \frac{3+\sqrt{3}}{6}$$

$$\text{for } k=2 \quad x_1^u = y_1^u = \frac{5-\sqrt{15}}{10}, \quad x_2^u = y_2^u = \frac{1}{2}$$

$$x_3^u = y_3^u = \frac{5+\sqrt{15}}{10}$$

The basis function $\lambda_{i,j}$ associated to the degree of freedom of coordinates (x_i, y_j) is given by :

$$\text{for } Q_{k+1}, \qquad \lambda_{i,j} = \frac{\displaystyle\prod_{\substack{l=1 \\ l \neq i}}^{k+2} (x - x_l) \prod_{\substack{l=1 \\ l \neq j}}^{k+2} (y - y_l)}{\displaystyle\prod_{\substack{l=1 \\ l \neq i}}^{k+2} (x_i - x_l) \prod_{\substack{l=1 \\ l \neq j}}^{k+2} (y_j - y_l)}$$

We can remark for example, in figure (6) that for $k = 1$ we have 3 types of basis functions associated to p_h,

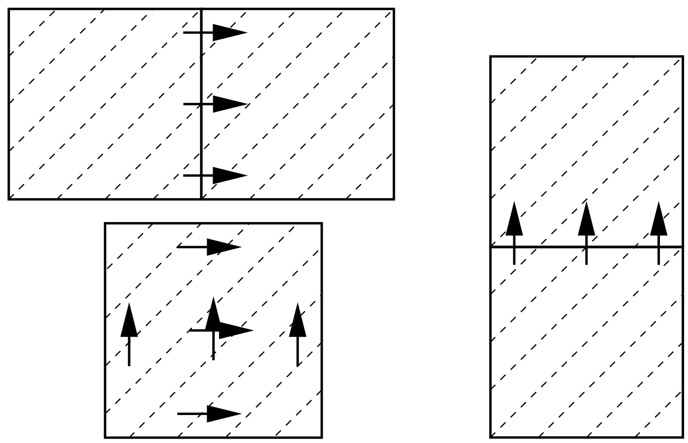

Figure 6: Support of Q_2 basis functions for p_h

and 1 type of basis functions for u_h (see figure 7)

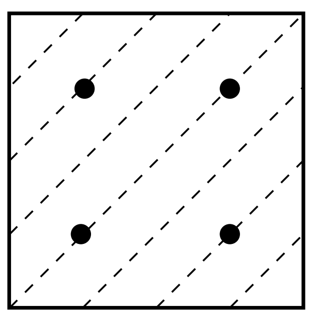

Figure 7: Support of Q_1 basis functions for u_h

Following the approach of P. Monk, G. Cohen and N. Tordjman (c.f [13], [20]), we approximate the integrals in (11(i), (ii)) by adequate quadrature formulas, for which the position of the quadrature points coincide with the position of the degrees of freedom. For our elements it consist in :

a) using the Gauss-Lobatto quadrature formulas for the approximation of the M_p matrix. The resulting matrix is now block diagonal. Each block is associated to one quadrature point and its dimension is equal to the number of degrees of freedom at this point (the "worst" case concerns the summits of the elements K, where the dimension of the local block is 4×4).

b) using the Gauss-Legendre quadrature formula for the approximation of the M_u matrix, the resulting matrix is diagonal.

Remark 2 *One can remark that M_h being chosen discontinuous, the matrix M_u does not really need to be mass lumped (it is always block diagonal).*

Remark 3 *The generalization of the previous techniques to higher orders ($k \geq 3$) can be done without great difficulties, using higher order Gauss Lobatto and Gauss Legendre quadrature formulas [9].*

Remark 4 *The Raviart Thomas k-order approximation consists in the choice*

$$\hat{X}^{RT} = RT_{[k]} = P_{k+1,k} \times P_{k,k+1} \quad \text{and} \quad \hat{M} = Q_k$$

It is clear that the new approximate space X_h contains the space X_h^{RT}. Thus we have enriched the p-approximation, while keeping the same v-approximate space.

3 Analysis of the new mixed finite element for an elliptic problem

Following [6], we will study in this section the mixed approximation of the elliptic problem which is in fact the stationary problem associated to the evolution problem (5). Actually, we give in paragraph 3.2 an abstract result for a class of elliptic problems posed in a more general framework and show in paragraph 3.3 that the model problem enters this framework. This analysis will be used in paragraph 3.4 to study an elliptic projection operator that will be useful for the analysis of the approximation of the evolution problem done in section 4.

3.1 The elliptic problem

The elliptic problem we consider here is :

$$(18) \qquad \left\{ \begin{array}{l} \text{Find } u \in H_0^1(\Omega) \text{ such that } : \\ \\ -\text{div } (A^{-1}(x)\nabla u) = f \qquad , \; f \in L^2(\Omega) \end{array} \right.$$

We know that (18) admits a unique solution $u \in H_0^1(\Omega)$ and there exists $c > 0$ such that :

$$(19) \qquad \|u\|_{H^1(\Omega)} \leq c \; \|f\|_{L^2(\Omega)}$$

As for the time dependent problem, we define

$$(20) \qquad p = A^{-1}(x)\nabla u$$

and this gives

$$(21) \qquad -\text{div } p = f$$

The mixed formulation associated to equations (20), (21) is :

$$(22) \qquad \left\{ \begin{array}{lll} \text{Find } (p,u) \in X \times M = H(\text{div };\Omega) \times L^2(\Omega) \text{ such that :} & & \\ \\ a(p,q) + b(u,q) & = 0 & \forall q \in X \\ \\ b(w,p) & = -(f,w) & \forall w \in M, \end{array} \right.$$

where $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are defined by (6) and satisfy properties (7). As it is proven in [7] there exist a unique solution (p,u) in $X \times M$ of problem (22) where u is also the solution of the initial problem (18) (in fact the abstract result yields the uniqueness of u only in $M/\text{ Ker } B^t$, but it is easy to check that for the divergence operator $KerB^t = \{0\}$). For the approximation of this problem, we again consider the finite dimensional spaces X_h and M_h defined by (8). The discrete problem associated to (22) is :

$$(23) \qquad \left\{ \begin{array}{lll} \text{Find } (p_h,u_h) \in X_h \times M_h \text{ such that :} & & \\ \\ a(p_h,q_h) + b(u_h,q_h) & = 0 & \forall q_h \in X_h \\ \\ b(w_h,p_h) & = -(f,w_h) & \forall w_h \in M_h, \end{array} \right.$$

The elliptic problem (22) and its approximation (23) have been studied by several authors, see for example [17], [7] and we know that it admits a unique solution (p_h, u_h) in $X_h \times M_h$ with the convergence property :

$$(p_h, u_h) \longrightarrow (p, u) \quad \in \quad X \times M \tag{24}$$

when the following assumptions are satisfied :

$$(25) \quad \left| \begin{array}{ll} (i) & \text{The uniform discrete inf-sup condition} \\ & \exists c > 0 \text{ independent of } h \text{ such that} \\ & \forall\, w_h \in M_h,\ \exists\, q_h \in X_h \,/\, b(w_h, q_h) \geq c\, \|w_h\|_M\, \|q_h\|_X \\ (ii) & \text{The coercivity of the form } a(\cdot,\cdot) \text{ on } V_h (= \text{Ker} B_h) \\ & \exists \alpha > 0 \text{ independent of } h \text{ such that} \\ & \forall\, p_h \in V_h = \{q_h \in X_h \,/\, b(w_h, q_h) = 0,\ \forall\, w_h \in M_h\}, \quad a(p_h, p_h) \geq \alpha\, \|p_h\|_X^2 . \end{array} \right.$$

These assumptions are satisfied by the lowest order Raviart Thomas element [17]. With our new choice for $\hat{X}$ it is easy to verify that property (25-i) is still true but we no more satisfy the second relation of (25). Indeed, consider for example the function $\vec{f_h}$:

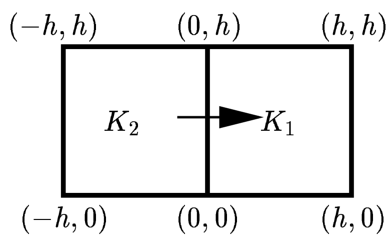

Figure 8: The function $\vec{f_h}$

we have

$$\vec{f_h}\,|_{K_1} = \begin{pmatrix} (1-\dfrac{x}{h})(1-2\dfrac{y}{h}) \\ 0 \end{pmatrix} \text{ and } \vec{f_h}\,|_{K_2} = \begin{pmatrix} (1+\dfrac{x}{h})(1-2\dfrac{y}{h}) \\ 0 \end{pmatrix}$$

We can easily see that $\vec{f_h} \in V_h$ and

$$a(f_h, f_h) \underset{h\to 0}{\longrightarrow} 0$$

$$\|f_h\|_X \underset{h\to 0}{\longrightarrow} \frac{2}{3}$$

so that we can not expect to verify (25-ii). We will prove although (see section 3) that this choice gives a good approximate solution and we will show a new convergence result.

Remark 5 *In order to preserve (25-ii) one could change the approximation space $\hat{M}$. The natural choice is :*

$$\hat{M} = P_1$$

the key point is that the divergence operator sends $Q_1 \times Q_1$ in P_1. We have eliminated this choice because it is rather expensive in terms of calculation time and memory requests.

3.2 An abstract result

The first point in our new theory is that we need to introduce a third Hilbert space. More precisely let M, X, H be three Hilbert spaces with :

$$X \subset H \quad \text{and} \quad |\cdot|_H \leq \|\cdot\|_X \tag{26}$$

The reader can have in mind that, for our application we shall have :

$$H = (L^2(\Omega))^2, \; X = H(\text{div}, \Omega) \text{ and } M = L^2(\Omega)$$

Take $a(\cdot,\cdot)$ et $b(\cdot,\cdot)$ two continuous bilinear forms in $H \times H$ and $M \times X$ verifying :

$$\left|\begin{array}{ll} (i) & \exists\, c > 0 \,/\, \forall\, w \in M, \; \exists\, q \in X \,/\, b(w,q) \geq c \, \|w\|_{M/\text{Ker } B^t} \, \|q\|_X \\ (ii) & \exists\, \alpha > 0 \,/\, \forall\, p \in V = \{q \in X \,/\, b(w,q) = 0, \; \forall w \in M\}, \quad a(p,p) \geq \alpha \, \|p\|_X^2 \end{array}\right. \tag{27}$$

From the continuity of the bilinear form $a(\cdot,\cdot)$ we know that there exists an operator $\mathcal{A}$ in $\mathcal{L}(H)$ such that :

$$a(p,q) = (\mathcal{A}p, q)_H \quad \forall p, q \; \in \; H \times H$$

We can also define operators $B : X \to M'$ and $B^t : M \to X'$ such that :

$$\langle Bp, w \rangle_{M' \times M} = \langle p, B^t w \rangle_{X \times X'} = b(w,p) \;\; \forall (p,w) \in X \times M$$

we then define the kernel of B and B^t as follows :

$$\text{Ker } B = \{p \in X \,/\, b(w,p) = 0, \; \forall w \in M\} \, (= V)$$

$$\text{Ker } B^t = \{w \in M \,/\, b(w,p) = 0, \; \forall p \in X\}$$

We shall identify the quotient space $M/\text{Ker } B^t$ with the orthogonal complement of Ker B^t

$$M/\text{Ker } B^t \equiv (Ker B^t)^{\perp} \equiv \{w \in M \,/\, (v,w)_M = 0, \forall\, v \in M\}$$

We are interested in the numerical approximation of the solution (p,u) of the following problem :

$$\left\{\begin{array}{lll} \text{Find } (p,u) \in X \times M \text{ such that :} & & \\ a(p,q) + b(u,q) & = 0 & \forall q \in X \\ b(w,p) & = -\langle f, w \rangle & \forall w \in M. \end{array}\right. \tag{28}$$

with $f \in M'$, the dual space of M. Under these assumptions, we have the classical result (see [7]) :

Theorem 1 *For all $f \in ImB$, problem (28) has a unique solution (p,u) in $X \times (M/\text{Ker } B^t)$. Moreover, one has the bound*

$$\|u\|_{M/\text{Ker } B^t} + \|p\|_X \leq C \, \|f\|_{M'}$$

Suppose $X_h \subset X$ and $M_h \subset M$ finite dimension approximation spaces. We consider then the approximate problem :

$$\left\{\begin{array}{lll} \text{Find } (p_h,u_h) \in X_h \times M_h \text{ such that} & & \\ a(p_h,q_h) + b(u_h,q_h) & = 0 & \forall q_h \in X_h \\ b(w_h,p_h) & = -\langle f, w_h \rangle & \forall w_h \in M_h. \end{array}\right. \tag{29}$$

We set :

$$V_h(f) = \{q_h \in X_h \,/\, b(w_h,q_h) = -(f,w_h), \; \forall w_h \in M_h\}$$

$$V_h = V_h(0) = \text{Ker } B_h$$

we make the following hypothesis :

(H0) $\forall f \in \mathrm{Im}\, B,\; V_h(f) \neq \emptyset$

(H1) Orthogonal decomposition of X_h:

$$\left| \begin{array}{l} X_h = X_h^s \oplus X_h^r \;\; (p_h = p_h^s + p_h^r) \;\;,\;\; X_h^r \subset V_h \\ \forall (p_h^s, p_h^r) \in X_h^s \times X_h^r \;\; (p_h^s, p_h^r)_H = 0 \end{array} \right.$$

(H2) "Strong" discrete uniform inf-sup condition :

there exists a constant $c > 0$, independent of h, such that

$$\forall\, w_h \in M_h, \;\; \exists\, q_h^s \in X_h^s \quad / \quad b(w_h, q_h^s) \geq c \, \|w_h\|_{M/\mathrm{Ker}\; B_h^t} \, \|q_h^s\|_X$$

(H3) "Weak" coercivity :

there exists a constant $\alpha > 0$, independent of h, such that

$$\forall\, p_h \in V_h, \;\; a(p_h, p_h) \geq \alpha \left(\|p_h^s\|_X^2 + |p_h^r|_H^2 \right)$$

(H4) Approximation properties :

$$\left\{ \begin{array}{ll} \lim\limits_{h \to 0} \inf\limits_{q_h^s \in X_h^s} \|p - q_h^s\|_X = 0 & \forall p \in X \\ \lim\limits_{h \to 0} \inf\limits_{w_h \in M_h} \|u - w_h\|_M = 0 & \forall u \in M. \end{array} \right.$$

We can introduce as in the case of the continuous problem operator B_h from X_h to M_h defined from :

$$(B_h p_h, w_h)_{M' \times M} = b(p_h, w_h) \;\; \forall p_h \in X_h, \; \forall w_h \in M_h$$

$$\left(p_h, B_h^t w_h\right)_{X \times X'} = b(p_h, w_h) \;\; \forall p_h \in X_h, \; \forall w_h \in M_h$$

We then define :

$$\mathrm{Ker} B_h = \{p_h \in X_h \,/\, b(p_h, w_h) = 0, \; \forall w_h \in M_h\} = V_h$$

$$\mathrm{Ker} B_h^t = \{w_h \in M_h \,/\, b(p_h, w_h) = 0, \; \forall p_h \in X_h\}$$

Remark 6 *It may be more convenient to characterize hypothesis (H0) with one of the following equivalent statement :*

$$(30) \qquad \begin{array}{ll} (H0) - (ii) & \forall p \in X, \;\; \exists p_h \in X_h \text{ such that } b(p - p_h, w_h) = 0, \;\; \forall w_h \in M_h \\ (H0) - (iii) & \mathrm{Ker}\; B_h^t = \mathrm{Ker}\; B^t \cap M_h \subset \mathrm{Ker}\; B^t \end{array}$$

Remark 7 *We call hypothesis (H2) "strong" as we suppose the existence of a q_h^s in X_h^s instead of X_h, which would be the classical assumption, more precisely we have $X_h^s \subsetneq X_h$. On the contrary hypothesis (H3) is "weaker" in the sense that forall $p_h \in X_h^s \cap V_h$ or $p_h \in X_h^r \cap V_h \equiv X_h^r$,*

$$\|p_h^s\|_X^2 + |p_h^r|_H^2 \leq \|p_h\|_X^2$$

In the abstract case, this condition is not clearly weaker for all $p_h \in V_h$. In our application, however, it is easy to check that this would be still true for any $p_h \in V_h$.

Theorem 2 *With the hypothesis (H0) to (H4), problem (29) admits a unique solution:*

$$(p_h = p_h^s + p_h^r, u_h) \in X_h \times (M_h/\text{Ker } B_h^t)$$

and we have the following convergence result :

- $(p_h^s, u_h) \quad \rightarrow (p,u) \quad$ *in* $X \times M$
- $p_h^r \quad \rightarrow 0 \quad$ *in* H.

More precisely, we obtain the error estimates :

$$|p_h^r|_H + \|p - p_h^s\|_X + \|u - u_h\|_{M/\text{Ker } B_h^t} \leq C\left\{\inf_{q_h^s \in X_h^s} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M + \inf_{z_h^s \in X_h^s} \|\mathcal{A}p - z_h^s\|_H\right\}.$$

Proof: First note that hypothesis (H0) and non-uniform discrete coercivity on the kernel Ker B_h, i.e.,

$$\exists \alpha_h > 0,\ \forall p_h \in V_h \ \ a(p_h, p_h) \geq \alpha_h \|p_h\|_X^2 \tag{31}$$

ensure existence and uniqueness of the solution (p_h, u_h) of the discrete problem (29) in $X_h \times (M_h/\text{Ker } B_h^t)$. Since in finite dimension all the norms are equivalent, the non-uniform discrete coercivity on the kernel Ker B_h (31) is a consequence of (H3).

Take q_h any element of $V_h(f)$, we also have $p_h \in V_h(f)$ (second equation of (29)) thus :

$$(p_h - q_h) \in V_h$$

We can write

$$a(p_h - q_h, p_h - q_h) = a(p - q_h, p_h - q_h) + a(p_h - p, p_h - q_h) \tag{32}$$

Taking the difference between the first equation of the continuous(28) and the discrete(29) problem we have:

$$a(p_h - p, p_h - q_h) = b(u - w_h, p_h - q_h) \quad \forall w_h \in M_h. \tag{33}$$

Using (H1) in (33), we can write, with obvious notations :

$$a(p_h - p, p_h - q_h) = b(u - w_h, p_h^s - q_h^s) + b(u, p_h^r - q_h^r) \tag{34}$$

(we have used the fact that $b(w_h, p_h^r - q_h^r) = 0$, since $X_h^r \subset V_h$), then by (28) we get :

$$\begin{aligned} a(p_h - p, p_h - q_h) &= b(u - w_h, p_h^s - q_h^s) - a(p, p_h^r - q_h^r) \\ &= b(u - w_h, p_h^s - q_h^s) - (\mathcal{A}p, p_h^r - q_h^r)_H \end{aligned} \tag{35}$$

We chose now :

$$\left| \begin{array}{l} q_h = q_h^s \in V_h^s(f) = V_h(f) \cap X_h^s \\ \\ q_h^r = 0 \end{array} \right. \tag{36}$$

from (32) through (35) and (36) we obtain

$$a(p_h - q_h^s, p_h - q_h^s) = a(p - q_h^s, p_h - q_h^s) + b(u - w_h, p_h^s - q_h^s) - (\mathcal{A}p, p_h^r)$$

or by using the orthogonality of X_h^r and X_h^s (H1)

$$a(p_h - q_h^s, p_h - q_h^s) = a(p - q_h^s, p_h - q_h^s) + b(u - w_h, p_h^s - q_h^s) - (\mathcal{A}p - z_h^s, p_h^r)_H \quad \forall\ z_h^s \in X_h^s \tag{37}$$

Further, from inequality

$$|p_h - q_h^s|_H \leq \|p_h^s - q_h^s\|_X + |p_h^r|_H$$

and (H1), (37) leads to :

$$\left| \begin{array}{ll} \alpha(\|p_h^s - q_h^s\|_X^2 + |p_h^r|_H^2) & \leq \|a\| \, |p - q_h^s|_H \, (\|p_h^s - q_h^s\|_X + |p_h^r|_H) \\ & + \|b\| \, \|u - w_h\|_M \, \|p_h^s - q_h^s\|_X + |\mathcal{A}p - z_h^s|_H \, |p_h^r|_H \, . \end{array} \right.$$

We deduce the existence of a constance C depending on $f,\ p,\ \|a\|\,,\ \|b\|\,,$ and α such that

$$\|p_h^s - q_h^s\|_X + |p_h^r|_H \leq C\,(|p - q_h^s|_H + \|u - w_h\|_M + |\mathcal{A}p - z_h^s|_H) \quad \forall (q_h^s, w_h, z_h^s) \in V_h^s(f) \times M_h \times X_h^s.$$

which gives

$$\left| \begin{array}{ll} \|p - p_h^s\|_X & \leq (1 + C) \left(\inf_{q_h^s \in V_h^s(f)} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M + \inf_{z_h^s \in X_h^s} |\mathcal{A}p - z_h^s|_H \right) \\ |p_h^r|_H & \leq C \left(\inf_{q_h^s \in V_h^s(f)} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M + \inf_{z_h^s \in X_h^s} |\mathcal{A}p - z_h^s|_H \right) . \end{array} \right.$$

To conclude let us recall that the inf-sup condition (H2) implies (cf. [7]):

$$\inf_{q_h^s \in V_h^s(f)} \|p - q_h^s\|_X \leq c_1 \inf_{q_h^s \in X_h^s} \|p - q_h^s\|_X \, .$$

Finally, it remains to prove estimates for $\|u - u_h\|_{M/\text{Ker } B_h^t}$. Let us subtract the first equation of (29) from the first equation of (28). We get

$$a(p - p_h, q_h) + b(u - u_h, q_h) = 0 \quad \forall q_h \in X_h$$

so that for any w_h in M_h it comes

$$b(u_h - w_h, q_h) = a(p - p_h, q_h) + b(u - w_h, q_h) \ \ \forall q_h \in X_h$$

choosing now $q_h = q_h^s$ implies

$$b(u_h - w_h, q_h^s) = a(p - p_h, q_h^s) + b(u - w_h, q_h^s) \ \ \forall q_h^s \in X_h^s$$

Using this and the inf-sup condition we have

$$\begin{array}{ll} \|u_h - w_h\|_{M/\text{Ker } B_h^t} & \leq \dfrac{1}{c} \sup_{q_h^s \in X_h^s} \dfrac{b(u_h - w_h, q_h^s)}{\|q_h^s\|_X} \\ & \leq \dfrac{1}{c} \sup_{q_h^s \in X_h^s} \dfrac{a(p - p_h, q_h^s) + b(u - w_h, q_h^s)}{\|q_h^s\|_X} \\ & \leq \dfrac{1}{c} \left(\|a\| \, \|p - p_h\|_H + \|b\| \, \|u - w_h\|_M \right) \end{array}$$

it follows from the triangle inequality

$$\|u - u_h\|_{M/\text{Ker } B_h^t} \leq C' \left\{ \inf_{w_h \in M_h} \|u - w_h\|_M + |p_h^r|_H + \|p - p_h^s\|_H \right\} . \square$$

Remark 8 *One could easily remark that the space X_h^s satisfies all assumptions (H0) to (H4), why not replacing then $X_h = X_h^s$? In fact we should imagine the case, in which whether we can not characterize the space X_h^s or we prefer the use of X_h (that is in particular our case for the evolution problem, where we prefer the use of X_h in order to achieve mass lumping).*

3.3 Application to the model problem

We have in the case of problem (22)

- $H = (L^2(\Omega))^2$
- $X = H(\text{div}\,, \Omega)$
- $M = L^2(\Omega)$

The operator B is in this case the divergence operator which is surjective from X into M which means that Ker $B^t = 0$. From remark 6 it follows that hypothesis (H0) is equivalent to the surjectivity of B_h from X_h into M_h. This is true for the classical Raviart Thomas approximations [17], i.e., B_h is surjective from X_h^{RT} into M_h. Since the new space X_h contains the space X_h^{RT}, as noticed in remark 4, it is straightforward to check (H0), therefore Ker $B_h^t = 0$.

3.3.1 Approximation with the lowest order finite element $Q_1^{div} - Q_0$

Let us begin by checking the hypothesis (H1). We take as approximation spaces :

$$\left\{ \begin{array}{l} X_h = \{q_h \in X \;/\; \forall K \in \mathcal{T}_h,\; q_h|_K \in Q_1 \times Q_1\} \\ \\ M_h = \{w_h \in M \;/\; \forall K \in \mathcal{T}_h,\; w_h|_K \in Q_0\} \end{array} \right.$$

We define now X_h^s as the lowest order Raviart-Thomas element (cf. [17]) $RT_{[0]}$:

$$X_h^s = \left\{ w_h \in X \;/\; \forall K \in \mathcal{T}_h,\; w_h \in RT_{[0]} \right\} \subset X.$$

In order to describe its orthogonal X_h^r, we denote as illustrated in figure 9

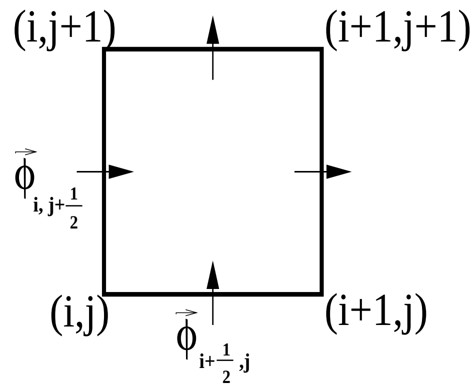

Figure 9: The $RT_{[0]}$ element

- (i, j) the node of $\mathcal{T}_h$ with coordinates (ih, jh),
- $(i + \dfrac{1}{2}, j)$ the horizontal side joining the nodes (i, j) and $(i + 1, j)$,
- $(i, j + \dfrac{1}{2})$ the vertical side joining the nodes (i, j) and $(i, j + 1)$.

The base functions $\vec{\phi}_{i,j+\frac{1}{2}}$, $\vec{\phi}_{i+\frac{1}{2},j} \in RT_{[0]}$ can be written as:

$$\vec{\phi}_{i,j+\frac{1}{2}} = \begin{pmatrix} \phi_{i,j+\frac{1}{2}} \\ \\ 0 \end{pmatrix}, \; \vec{\phi}_{i+\frac{1}{2},j} = \begin{pmatrix} 0 \\ \\ \phi_{i+\frac{1}{2},j} \end{pmatrix}$$

where $\phi_{i,j+\frac{1}{2}} \in P_{10}$ and $\phi_{i+\frac{1}{2},j} \in P_{01}$ and defined by :

$$\phi_{i,j+\frac{1}{2}} = \frac{(x_{i+1} - x)}{h}, \; \phi_{i+\frac{1}{2},j} = \frac{(y_{i+1} - y)}{h}$$

It is then easy to prove the following Lemma :

Lemma 1 *The space $X_h^r = (X_h^s)^{\perp}$ can be generated from the functions :*

$$\vec{\psi}_{i+\frac{1}{2},j} = \begin{pmatrix} \left(x - x_{i+\frac{1}{2}}\right) \phi_{i+\frac{1}{2},j} \\ \\ 0 \end{pmatrix} \; \textit{and} \; \vec{\psi}_{i,j+\frac{1}{2}} = \begin{pmatrix} 0 \\ \\ \left(y - y_{j+\frac{1}{2}}\right) \phi_{i,j+\frac{1}{2}} \end{pmatrix}$$

where $x_{i+\frac{1}{2}} = (i + \frac{1}{2})h$ and $y_{j+\frac{1}{2}} = (j + \frac{1}{2})h$. Moreover $X_h^r \subset V_h$.

Using now the well known properties of $X_h^s - M_h$ (cf. [17], [18], [7]) hypothesis (H2) and (H4) are trivial, only (H3) needs to be checked.

Lemma 2 *For all p_h in V_h, we have :*

$$a(p_h,p_h) \geq \alpha \left(\|p_h^s\|_X^2 + |p_h^r|_H^2 \right)$$

α being defined from (1).

Proof: From (1) we get:

$$a(p_h,p_h) \geq \alpha \, \|p_h\|_0^2 = \alpha \left(\|p_h^s\|_0^2 + \|p_h^r\|_0^2 \right)$$

Remark now that $p_h \in V_h$ implies $p_h^s \in V_h \cap X_h^s$. Hence it follows (cf. [7]):

$$\begin{cases} (i) & \text{div } p_h^s \in M_h \\ (ii) & (\text{ div } p_h^s, w_h) = 0 \quad \forall w_h \in M_h. \end{cases}$$

Recalling (i) we can take

$$w_h = \text{ div } p_h^s$$

in (ii), which gives:

$$\text{div } p_h^s = 0.$$

The claim is thus proved.□

We can apply now Theorem 2 to the approximation problem (23), and by using the usual interpolation results (cf. [18]) we obtain : (here we have Ker $B_h^t = 0$)

Theorem 3 *The problem (23) admits a unique solution:*

$$(p_h, u_h) \in X_h \times M_h.$$

which satisfies:

- $(p_h^s, u_h) \quad \to (p,u) \quad$ *in* $H(\text{div }, \Omega) \times L^2(\Omega)$
- $p_h^r \quad \to 0 \quad$ *in* $L^2(\Omega)$.

Further if we assume the solution to be more regular, $(p,u) \in H^1(div,\Omega) \times H^1(\Omega)$, we get:

$$\|p_h^r\|_{L^2} + \|p - p_h^s\|_{H(\text{div })} + \|u - u_h\|_{L^2} \leq Ch \left(|u|_{H^1} + |p|_{H^1} + |div\ p|_{H^1} + |\mathcal{A}p|_{H^1} \right)$$

where $|.|_{H^1}$ denotes the usual semi norm in $H^1(\Omega)$.

Remark 9 *Let us now consider the case of isotropic waves equation, which corresponds in taking $A(x)$ in (18) as a diagonal matrix($A_d(x)$). It is then easy to prove the following : the approximate problem*

$$(38) \qquad \begin{cases} \textit{Find } (p_h, u_h) \in X_h \times M_h \textit{ such that} \\ a_d(p_h, q_h) + b(u_h, q_h) = 0 & \forall q_h \in X_h \\ b(w_h, p_h) = -\langle f, w_h \rangle & \forall w_h \in M_h. \end{cases}$$

admits a unique solution $(p_h = p_h^s + p_h^r, u_h)$ with $p_h^r \equiv 0$ and (p_h^s, u_h) the solution of the following problem :

$$\begin{cases} \textit{Find } (p_h^s, u_h) \in X_h^s \times M_h \textit{ such that} \\ a_d(p_h^s, q_h^s) + b(u_h, q_h^s) = 0 & \forall q_h^s \in X_h^s \\ b(w_h, p_h^s) = -\langle f, w_h \rangle & \forall w_h \in M_h. \end{cases}$$

where $a_d(\cdot,\cdot)$ *is given by (6) after replacing* $A(x)$ *by* $A_d(x)$.

Noting that in this particular case we have :

$$a_d(p_h^r, q_h^s) = 0 \quad \forall q_h^s \in X_h^s$$

$$a_d(p_h^s, q_h^r) = 0 \quad \forall q_h^r \in X_h^r$$

and using the fact that

$$X_h^r \subset V_h$$

we can decompose (38) in two independent problems in X_h^s *and* X_h^r *:*

$$\left\{\begin{array}{lll} \text{Find } (p_h^s, u_h) \in X_h^s \times M_h \text{ such that} & & \\ a_d(p_h^s, q_h^s) + b(u_h, q_h^s) & = 0 & \forall q_h^s \in X_h^s \\ b(w_h, p_h^s) & = -\langle f, w_h\rangle & \forall w_h \in M_h. \end{array}\right.$$

and

$$(39) \qquad \left\{\begin{array}{ll} \text{Find } (p_h^r) \in X_h^r \text{ such that} & \\ a_d(p_h^r, q_h^r) \quad = 0 & \forall q_h^r \in X_h^r \end{array}\right.$$

It is obvious then from (39) that

$$p_h^r \equiv 0$$

This remark is no longer true when $A(x)$ *is not diagonal.*

3.3.2 Approximation with higher order finite elements, $Q_{k+1}^{div} - Q_k$

We have seen in section 2.3 that the natural generalization of the lowest order element consists in taking :

$$\hat{X} = X_k = Q_{k+1} \times Q_{k+1} \quad \text{and} \quad \hat{M} = M_k = Q_k$$

which leads us to introduce the spaces :

$$(40) \qquad (17) \quad \left\{\begin{array}{l} X_h = \{q_h \in X \;/\; \forall K \in \mathcal{T}_h,\; q_h|_K \in X_k\} \\ M_h = \{w_h \in M \;/\; \forall K \in \mathcal{T}_h,\; w_h|_K \in M_k\} \end{array}\right.$$

We are going to show that we can apply the abstract theory of section 3.2 to these spaces, the main difficulty lying in the construction of an orthogonal decomposition of X_h such that assumptions (H0) to (H4) are satisfied. In fact we shall deduce such a global decomposition from a local decomposition of the space $\hat{X}$, considered as a subspace of $L^2(K)$, where K is a single element.First we recall the definition of the Raviart-Thomas space $RT_{[k]}$ as :

$$RT_{[k]} = P_{k+1,k} \times P_{k,k+1}$$

which obviously satisfies

$$RT_{[k]} \subset X_k$$

Let us define $\Psi_k(K)$ as the orthogonal complement in X_k of $RT_{[k]}$ (for the inner product $L^2(K)$):

$$\Psi_k(K) = \left\{\psi \in \hat{X} \;/\; \int_K \psi\phi dx = 0,\; \forall \phi \in RT_{[k]}\right\}$$

Note that :

$$\dim \Psi_k(K) = 2(k+2)$$

The main property of the space $\hat{\Psi}$ is the following :

Lemma 3 *For any ψ in Ψ_k and v in M_k, one has :*

$$\int_K div\ \psi \cdot v\ dx = 0 \tag{41}$$

Proof :
For simplicity we shall take in the proof the reference element $K = [0,1] \times [0,1]$.
(i) We begin by a characterization of $\hat{\Psi}$. Let us introduce σ_k as the polynomial of one variable of degree $k+1$ such that :

$$P_{k+1} = P_k \oplus [\sigma_k]$$

Equivalently, σ_k is defined, up to multiplicative constant by :

$$\left\{ \begin{array}{l} \displaystyle\int_0^1 \sigma_k(x)p(x)dx = 0\ \forall p \in P_k \\ \sigma_k \in P_{k+1},\ \sigma_k \neq 0 \end{array} \right.$$

Then we claim that :

$$\hat{\Psi} = \tilde{\Psi} \equiv \left\{ \begin{pmatrix} p_1(x_1)\sigma_k(x_2) \\ \\ p_2(x_2)\sigma_k(x_1) \end{pmatrix},\ (p_1,p_2) \in P_{k+1} \times P_{k+1} \right\}$$

Indeed, as $\dim \tilde{\Psi} = 2(k+2)$, it is sufficient to check that $\tilde{\Psi}$ is orthogonal to $RT_{[k]}$. To see that, we first remark that (see [17]) $RT_{[k]}$ is generated by vector fields of the form

$$\left\{ \phi(x_1,x_2) = \begin{bmatrix} \tau_1(x_1)q_1(x_2) \\ \\ \tau_2(x_2)q_2(x_1) \end{bmatrix} \quad \text{with} \quad (\tau_1,\tau_2) \in P_{k+1} \times P_{k+1},\ (q_1,q_2) \in P_k \times P_k \right.$$

Let us consider

$$\left\{ \psi(x_1,x_2) = \begin{bmatrix} p_1(x_1)\sigma_k(x_2) \\ \\ p_2(x_2)\sigma_k(x_1) \end{bmatrix} \quad \text{with} \quad (p_1,p_2) \in P_{k+1} \times P_{k+1} \right. \tag{42}$$

we have :

$$\begin{array}{rcl} (\phi,\psi)_{L^2(K)} & = & \left(\displaystyle\int_0^1 \tau_1(x_1)p_1(x_1)dx_1 \right) \left(\displaystyle\int_0^1 \sigma_k(x_2)q_1(x_2)dx_2 \right) \\ & + & \left(\displaystyle\int_0^1 \sigma_k(x_1)q_2(x_1)dx_1 \right) \left(\displaystyle\int_0^1 \tau_2(x_2)p_2(x_2)dx_2 \right) \end{array}$$

As (q_1,q_2) are in P_k, by definition of σ_k, we obtain

$$\int_0^1 \sigma_k(x_2)q_1(x_2)dx_2 = \int_0^1 \sigma_k(x_1)q_2(x_1)dx_1 = 0 \Rightarrow (\phi,\psi)_{L^2(K)} = 0$$

(ii) To prove (41) we first use Green's formula :

$$\int_0^1 \mathrm{div}\psi v dx = -\int_0^1 \psi \cdot \nabla v dx + \int_{\partial K} (\psi \cdot n) v d\gamma$$

If v belongs to Q_{k+1}, ∇v belongs to $RT_{[k]}$. There for, since ψ belongs to $\hat{\Psi}$:

$$\int_0^1 \mathrm{div}\psi\ v\ dx = \int_{\partial K} (\psi \cdot n) v d\gamma$$

Let us decompose ∂K as :

$$\partial K = T_1 \cup T_2 \cup T_3 \cup T_4$$

according to figure 10.

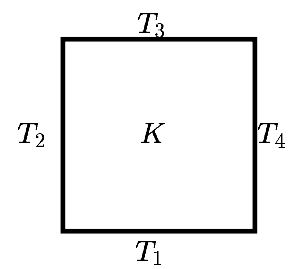

Figure 10: The edges of K.

Then we check that :

$$\forall 1 \leq j \leq 4,\ \forall q \in R_k(T_j) \int_{T_j} (\psi \cdot n) q d\gamma = 0$$

where $R_k(T_j)$ is the set of polynomials of degree k with respect to the abcisse along T_j. Let us consider for instance $j = 1$, and assume that ψ is given by (42), then

$$\int_{T_1} (\psi \cdot n) q d\gamma = -\left(\int_0^1 \sigma_k(x_1) q(x_1) \right) p_2(0) \equiv 0$$

To conclude, it suffices to remark that if v belongs to Q_{k+1} then $v|_{T_j}$ belongs to $R_k(T_j)$. This result suggest to define :

(43) $$X_h^s = \left\{ p_h \in H(\mathrm{div};\Omega) \;/\; \forall K \in \mathcal{T}_h,\ p_h|_K \in RT_{[k]} \right\}$$

and

$$X_h^r = \{ p_h \in H(\mathrm{div};\Omega) \;/\; \forall K \in \mathcal{T}_h,\ p_h|_K \in \Psi_k \}$$

We can now show the main result of this section

Theorem 4 *One has the decomposition*

$$X_h = X_h^s \oplus X_h^r$$

and assumptions (H0) to (H4) are satisfied.

Proof: (H0), (H3) and (H4) are classical properties of the Raviart-Thomas approximation spaces [17]. (H1) and (H2) are straightforward consequences (decomposing the integrals over Ω as the sum of the integrals over elements K) of the definition of Ψ_k and Lemma 3.

Remark 10 *For $k = 0$ we obtain the orthogonal decomposition of X_h described in section 3.3.1 that we resume in figure 11.*

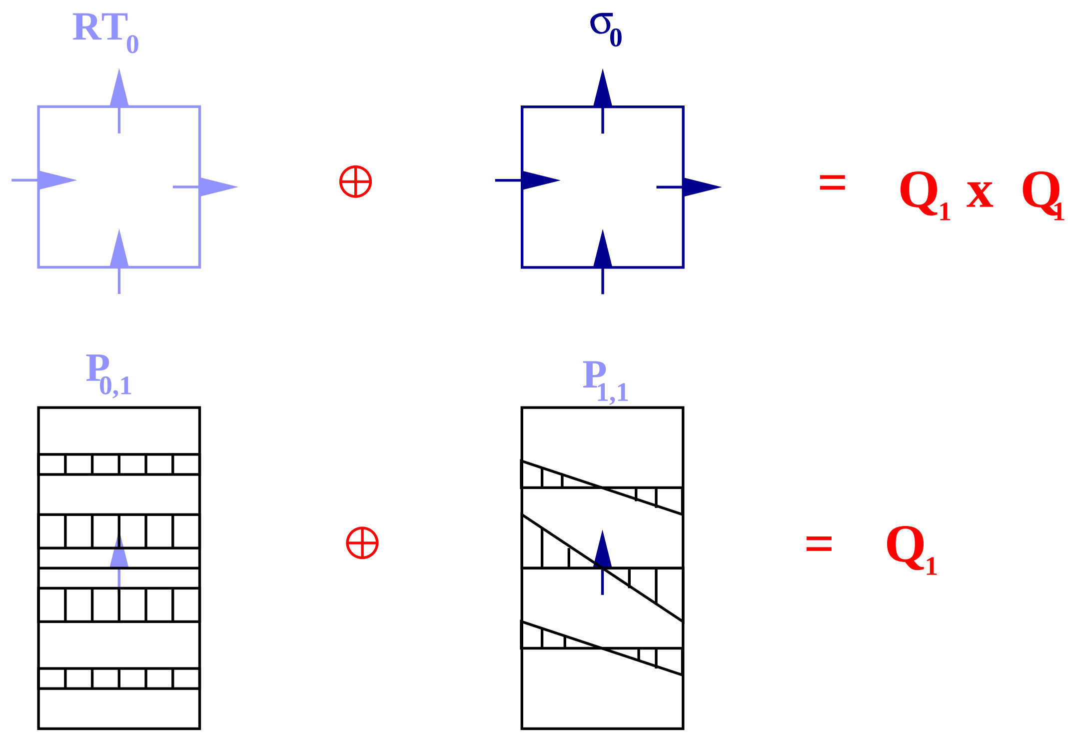

Figure 11: Orthogonal decomposition of $Q_1 \times Q_1$

In the same way we can resume the orthogonal decomposition of X_h for any k in figure 12

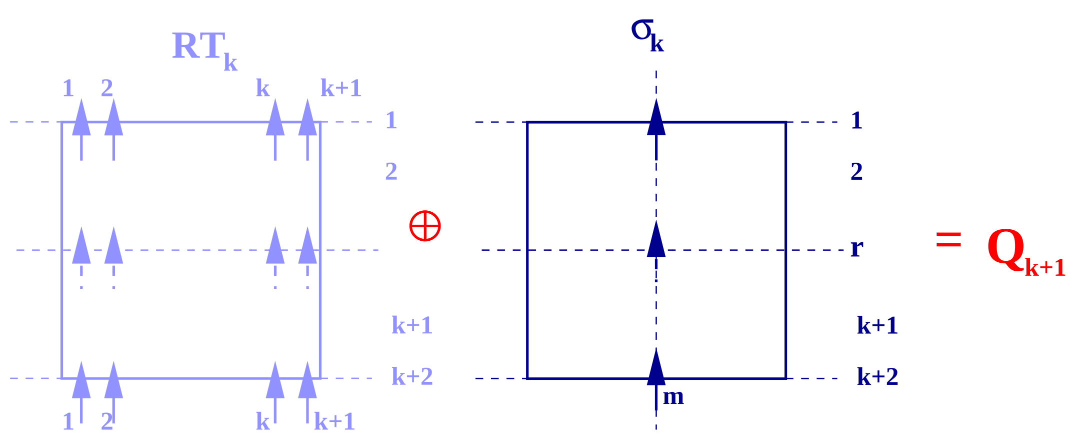

Figure 12: Orthogonal decomposition of Q_{k+1}

We can apply now Theorem 1 to the approximation problem (23), and by using the usual interpolation results (cf. [18]) in $X_h^s - M_h$ defined by (17), (43) we obtain :

Theorem 5 *The problem (23) admits a unique solution:*

$$(p_h, u_h) \in X_h \times M_h.$$

which satisfies:

- $(p_h^s, u_h) \quad \to (p, u) \quad$ *in* $H(\mathrm{div}\ , \Omega) \times L^2(\Omega)$
- $p_h^r \quad \to 0 \quad$ *in* $L^2(\Omega)$.

Further if we assume that the solution (p,u) of (22) is such that $(p,u) \in (H^m(\Omega))^2 \times (H^m(\Omega))$ and $div\, p \in H^m(\Omega)$ for $m = k+1$, then we get the error estimate :

$$\|p_h^r\|_{L^2} + \|p - p_h^s\|_{H(\text{div})} + \|u - u_h\|_{L^2} \leq C\ h^m \left(|u|_{H^m} + |p|_{H^m} + |div\ p|_{H^m} + |\mathcal{A}p|_{H^m}\right) \tag{44}$$

where $|.|_{H^m}$ denotes the usual semi norm in $H^m(\Omega)$.

Remark 11 *In order to expect the error bound (44), one has to suppose that the coefficients A_{ij} of A are sufficiently regular, for instance $A_{ij} \in C^m$.*

3.4 Application to the elliptic projection operator

We come back to the abstract framework described in section 3.2. We consider the problem : find $(\widehat{p}_h, \widehat{u}_h)$ such that

$$\left\{ \begin{array}{lll} a(p - \widehat{p}_h, q_h) + b(u - \widehat{u}_h, q_h) & = 0 & \forall q_h \in X_h \\ b(w_h, p - \widehat{p}_h) & = 0 & \forall w_h \in M_h, \end{array} \right. \tag{45}$$

We set

$$D(B^t) = \{w \in M \ /\ b(w,q) \leq C(w)\, \|q\|_H\,,\ \forall q \in X\} = \{w \in M\ /\ B^t w \in H\}$$

For $w \in D(B^t)$, we have

$$b(w,q) = (B^t w, q)_H,\ \ \forall\ q \in X$$

Let us introduce a notation :

$$\left\{ \begin{array}{l} |||p - \widehat{p}_h||| = \|p - \widehat{p}_h^{\,s}\|_X + \|\widehat{p}_h^{\,r}\|_H \\ \|\ [(p,u) - \Pi_h(p,u)]\ \| = |||p - \widehat{p}_h||| + \|u - \widehat{u}_h\|_M \end{array} \right. \tag{46}$$

It is straightforward, by application of the abstract result, to get the interpolation estimates :

Theorem 6 *We make hypothesis (H0) to (H4).*

(i) Forall $(p,u) \in X \times M$, problem (45) admits a unique solution

$$\Pi_h(p,u) = (\widehat{p}_h^{\,s} + \widehat{p}_h^{\,r}, \widehat{u}_h) \in X_h \times (M_h/\text{Ker } B_h^t)$$

(ii) Forall $(p,u) \in X \times D(B^t)$, there exists a constant C independent of h such that

$$\|\ [(p,u) - \Pi_h(p,u)]\ \| \leq C\ \mathcal{N}_h(p,u)$$

where

$$\mathcal{N}_h(p,u) = \inf_{q_h^s \in X_h^s} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M + \inf_{z_h^s \in X_h^s} \|B^t u - z_h^s\|_H$$

which shows in particular that $|||p - \widehat{p}_h|||$ and $\|u - \widehat{u}_h\|_M$ tend to zero.

Proof of theorem 6 : (i) The existence and uniqueness still comes from hypothesis (H0) and (H3) (with $f = Bp$ which is of course in $Im\ B$).

(ii) The proof of this point is a minor modification of the proof of theorem 2. The only difference concerns the treatment of term $b(u, p_h^r - q_h^r)$ in (34). In order to get the error estimate, it is necessary to relate this term to the H-scalar product. In the proof of theorem 2, this was done using the first continuous equation which implies that $b(u, p_h^r - q_h^r) = -a(p, p_h^r - q_h^r) \equiv -(\mathcal{A}p, p_h^r - q_h^r)_H$. Now we assume that $u \in D(B^t)$ so we have

$$b(u, p_h^r - q_h^r) = (B^t u, p_h^r - q_h^r)_H$$

The result is thus obtained by replacing $-\mathcal{A}p$ with $B^t u$. □

4 Error estimates for the evolution problem

Let us now come back to the initial evolution problem (5), (4) and see how we can relate the error estimates to the one obtained for the elliptic problem (22). Although we have constructed this element in order to be able to do mass lumping, we analyze the error for the discrete problem without mass lumping. Of course, when doing mass lumping, one should add to this error the quadrature error due to the numerical integration (see [20], [2]).

4.1 From error estimates for the elliptic problem to error estimates for the evolution problem

In this part, we use the same notations and hypothesis as in section 3.2, and we consider the evolution problem

$$(47)\qquad \begin{cases} \text{Find } (p,v):[0,T]\rightarrow X\times M \text{ such that :} \\ \dfrac{d}{dt}a(p,q)+b(v,q) = 0 \qquad \forall q\in X \\ \dfrac{d}{dt}(v,w)-b(w,p) = (f,w) \quad \forall w\in M, \\ p(0)=p_0 \ ; \ v(0)=v_0 \end{cases}$$

or equivalently in an operator form, if we assume enough regularity on the solution in time :

$$(48)\qquad \begin{cases} \text{Find } (p,v)\in (C^1(0,T;H)\cap C^0(0,T;X))\times C^1(0,T;M) \text{ such that :} \\ \mathcal{A}\dfrac{dp}{dt}+B^tv = 0 \quad \text{in } X' \\ \dfrac{dv}{dt}-Bp = f \quad \text{in } M', \\ p(0)=p_0 \ ; \ v(0)=v_0 \end{cases}$$

In the following, we use the notation $\mathcal{C}^{m,r}=C^m(0,T;H)\cap C^r(0,T;X)$.

Suppose $X_h\subset X$ and $M_h\subset M$ finite dimension approximation spaces. We consider then the approximate problem :

$$(49)\qquad \begin{cases} \text{Find } (p_h,v_h):[0,T]\rightarrow X_h\times M_h \text{ such that :} \\ \dfrac{d}{dt}a(p_h,q_h)+b(v_h,q_h) = 0 \qquad \forall q_h\in X_h \\ \dfrac{d}{dt}(v_h,w_h)-b(w_h,p_h) = (f,w_h) \quad \forall w_h\in M_h, \\ p_h(0)=p_{0,h} \ ; \ v_h(0)=v_{0,h} \end{cases}$$

From the classical theory of ODE's, we have the following result :

Theorem 7 *If $f\in C^0(0,T;M_h)$, then problem (49) has a unique solution $(p_h,v_h)\in C^1(0,T;X_h)\times C^1(0,T;M_h)$*

Moreover, we make hypotheses (H0) to (H4), in particular X_h admits the following orthogonal decomposition

$$X_h = X_h^s\oplus X_h^r$$

Following [8], [10] we introduce the elliptic operator defined in (45). By application of theorem 6, we get the following interpolation results :

Lemma 4 *Let (p,v) be the solution of (48) and assume that $(p,v)\in\mathcal{C}^{1,0}\times C^1(0,T;M)$, then*
(i) There exists a primitive of v, $u\in C^1(0,T;M)$, satisfying,

$$(50)\qquad \begin{cases} \dfrac{du}{dt}=v, \\ \mathcal{A}p_0+B^tu(0)=0 \end{cases}$$

This primitive is unique up to a constant element of Ker B^t.

(ii) $\forall t \in [0,T]$, *problem (45) admits a unique solution* $\Pi_h(p,u)(t) = (\widehat{p}_h, \widehat{u}_h)(t) \in X_h \times (M_h/\text{Ker } B_h^t)$ *and there exists a constant* C *independent of* h *such that*

$$\| [(p,u) - \Pi_h(p,u)]\,(t)\| \leq C\ \mathcal{N}_h(p,u)(t) \tag{51}$$

where $\mathcal{N}_h$ *is defined by,*

$$\mathcal{N}_h(p,u)(t) = \inf_{q_h^s \in X_h^s} \|p(t) - q_h^s\|_X + \inf_{w_h \in M_h} \|u(t) - w_h\|_M + \inf_{z_h^s \in X_h^s} \|\mathcal{A}p(t) - z_h^s\|_H$$

which shows in particular that $|||p - \widehat{p}_h|||$ *and* $\|u - \widehat{u}_h\|_M$ *tend to zero uniformly in time* ($t \in [0,T]$) ($\|\,[.]\,\|$ *and* $|||.|||$ *are defined in (46))* .

(iii) In the same way, if $(p,u) \in C^k(0,T;X) \times C^k(0,T;M)$, $k \geq 1$, *there exists a constant* C *independent of* h *such that*

$$\left\| \left[(\partial_t^k p, \partial_t^k u) - \Pi_h\left(\partial_t^k p, \partial_t^k u\right)\right](t)\right\| \leq C\ \mathcal{N}_h\left(\partial_t^k p, \partial_t^k u\right)(t) \tag{52}$$

where we used the notation $\partial_t^k g = \dfrac{d^k g}{dt^k}$

Remark 12 *Operators* Π_h *and* ∂_t^k *commute, and in particular we will set in the following :*

$$\widehat{v}_h = \partial_t\left(\widehat{u}_h\right) = \left(\widehat{\partial_t u}\right)_h \tag{53}$$

Proof of lemma 4 :

(i) We set $f_0 = -Bp_0 \in Im\ B$. From hypothesis (27), we know that there is a unique $(p_0, u_0) \in X \times (M/\text{Ker } B^t)$ such that

$$\begin{cases} a(p_0,q) + b(u_0,q) & = 0 \qquad \forall q \in X \\ b(w,p_0) & = -(f_0,w) \quad \forall w \in M, \end{cases}$$

which means that, p_0 being fixed, there is a unique $u_0 \in (M/\text{Ker } B^t)$ such that $Ap_0 + B^t u_0 = 0$. Now we define u as

$$u(t) = u_0 + \int_0^t v(s)ds$$

It is clear that $u \in C^1(0,T;M)$ and is the unique solution of (50).

(ii) Let $u \in C^1(0,T;M)$ be the primitive of v, plugging this into the first equation of (48) gives

$$\frac{d}{dt}(\mathcal{A}p + B^t u) = 0 \Longrightarrow (\mathcal{A}p + B^t u)(t) = \mathcal{A}p_0 + B^t u(0) = 0$$

thus $(p,u) \in C^1(0,T;X) \times C^1(0,T;M)$ satisfies

$$a(p,q) + b(u,q) = 0, \quad \forall q \in X$$

so we have $u \in D(B^t)$ and $B^t u = -\mathcal{A}p$. Applying theorem 6, we get the existence and uniqueness of the elliptic projection, for t fixed, and also the error estimate (51).

(iii) If (p,u) is sufficiently regular in time, we can differentiate with respect to t and get

$$\mathcal{A}\partial_t^k p + B^t \partial_t^k u = 0$$

We apply again theorem 6 to get (52).□

We now give the main result :

Theorem 8 *We make hypothesis (H0) to (H4). Let (p,v) be the solution of (48) and (p_h,v_h) the solution of the approximate problem (49) with the following initial conditions :*

$$(p_{0,h},v_{0,h}) = \Pi_h(p_0\ ,\ v_0) \tag{54}$$

- *If $(p,v)\in C^2(0,T;X)\times C^1(0,T;M)$, we have the following convergence result : $\forall t\in[0,T]$,*

$$\|p-p_h^s\|_H(t)\to 0\ ;\ \|p_h^r\|_H(t)\to 0\ ;\quad \|v-v_h\|_M(t)\to 0$$

More precisely, we obtain the error estimates :

$$\left\{\begin{aligned} &\|p-p_h^s\|_H(t)+\|p_h^r\|_H(t)\le C\left(\mathcal{N}_h(p,u)(t)+\int_0^t \mathcal{N}_h\left(\partial_t^2 p,\partial_t v\right)(s)ds\right)\\ &\|v-v_h\|_M(t)\le C\left(\mathcal{N}_h\left(\partial_t p,v\right)(t)+\int_0^t \mathcal{N}_h\left(\partial_t^2 p,\partial_t v\right)(s)ds\right)\end{aligned}\right. \tag{55}$$

- *In addition, if $(p,v)\in C^3(0,T;X)\times C^2(0,T;M)$, and $(p_h,v_h)\in C^2(0,T;X_h)\times C^2(0,T;M_h)$, we have the following convergence result in norm X: $\forall t\in[0,T]$, $\|p-p_h^s\|_X(t)\to 0$. More precisely :*

$$\|p-p_h^s\|_X(t)\le C\left(\mathcal{N}_h(p,u)(t)+\mathcal{N}_h\left(\partial_t^2 p,\partial_t v\right)(t)+\int_0^t\left(\mathcal{N}_h\left(\partial_t^2 p,\partial_t v\right)(s)+\mathcal{N}_h\left(\partial_t^3 p,\partial_t^2 v\right)(s)\right)ds\right) \tag{56}$$

In order to prove theorem 8 we need the following lemma :

Lemma 5 *Let (p,v) be the solution of (48) and (p_h,v_h) the solution of the approximate problem (49) with the initial conditions (54). Let $\Pi_h(p,u)=(\widehat{p}_h^{\,s}+\widehat{p}_h^{\,r},\widehat{u}_h)$ the elliptic projection defined in lemma 4, and $\widehat{v}_h$ defined in (53). We set $\delta_h=v-\widehat{v}_h$.*

(i) If $(p,v)\in\mathcal{C}^{1,0}\times C^1(0,T;M)$, there exists a constant C_1, independent of h such that $\forall t\in[0,T]$

$$\|\widehat{p}_h^{\,s}-p_h^s\|_H(t)+\|\widehat{p}_h^{\,r}-p_h^r\|_H(t)+\|\widehat{v}_h-v_h\|_M(t)\le C_1\int_0^t\|\partial_t\delta_h\|_M(s)\ ds \tag{57}$$

(ii) Moreover if $(p,v)\in\mathcal{C}^{2,1}\times C^2(0,T;M)$ and $(p_h,v_h)\in C^2(0,T;X_h)\times C^2(0,T;M_h)$, there exists a constant C_2, independent of h such that $\forall t\in[0,T]$

$$\|\widehat{p}_h^{\,s}-p_h^s\|_X(t)\le C_2\left\{\|\partial_t\delta_h\|_M(t)+\int_0^t\left(\|\partial_t\delta_h\|_M(s)+\|\partial_t^2\delta_h\|_M(s)\right)ds\right\} \tag{58}$$

Proof of Lemma 5 :

- Estimation (57) : we begin by rewriting equations (47) with the test functions $q=q_h\in X_h\subset X$ and $w=w_h\in M_h\subset M$ and we subtract it from (49) :

$$\left\{\begin{array}{lll} \dfrac{d}{dt}a(p-p_h,q_h)+b(v-v_h,q_h) & =0, & \forall q_h\in X_h\\ \dfrac{d}{dt}(v-v_h,w_h)-b(w_h,p-p_h) & =0, & \forall w_h\in M_h,\\ (p-p_h)(0)=p_0-p_{0,h}\ ; & (v-v_h)(0)=v_0-v_{0,h} & \end{array}\right.$$

Introducing the elliptic projection $\Pi_h(p,u)=(\widehat{p}_h,\widehat{u}_h)$, we split the error between the approximate solution and the exact solution into two parts :

$$\left\{\begin{aligned} (p-p_h)(t)=&\quad(p-\widehat{p}_h)(t)+(\widehat{p}_h-p_h)(t)\\ (v-v_h)(t)=&\quad(v-\widehat{v}_h)(t)+(\widehat{v}_h-v_h)(t)\end{aligned}\right. \tag{59}$$

and we choose as approximate initial conditions the elliptic projection of the exact initial condition, (54), so that at time $t=0$ we have

$$(\widehat{p}_h-p_h)(0)=0\ ;\ \widehat{v}_h-v_h(0)=0$$

Using now the error decomposition (59) we obtain :

$$(60)\qquad \left\{\begin{array}{ll} a(\partial_t(\widehat{p}_h - p_h), q_h) + b(\widehat{v}_h - v_h, q_h) & = -a(\partial_t(p-\widehat{p}_h), q_h) - b(v-\widehat{v}_h, q_h) \quad \forall q_h \in X_h \\ (\partial_t(\widehat{v}_h - v_h), w_h) - b(w_h, \widehat{p}_h - p_h) & = -(\partial_t(v-\widehat{v}_h), w_h) + b(w_h, p-\widehat{p}_h) \quad \forall w_h \in M_h, \end{array}\right.$$

Differentiating the first equation of (45) (written for (p,u)) with respect to t we see that

$$\left\{\begin{array}{ll} a(\partial_t(p-\widehat{p}_h), q_h) + b(v-\widehat{v}_h, q_h) & = 0 \quad \forall q_h \in X_h \\ b(w_h, p-\widehat{p}_h) & = 0 \quad \forall w_h \in M_h, \end{array}\right.$$

that we plug into (60) to get

$$(61)\qquad \left\{\begin{array}{lll} a(\partial_t(\widehat{p}_h - p_h), q_h) + b(\widehat{v}_h - v_h, q_h) & = 0 & \forall q_h \in X_h \\ (\partial_t(\widehat{v}_h - v_h), w_h) - b(w_h, \widehat{p}_h - p_h) & = -(\partial_t(v-\widehat{v}_h), w_h) & \forall w_h \in M_h, \end{array}\right.$$

Further by taking $q_h = \widehat{p}_h - p_h,\ w_h = \widehat{v}_h - v_h$ in (60) and by adding the two equations, we get :

$$(62)\qquad a(\partial_t(\widehat{p}_h - p_h), \widehat{p}_h - p_h) + (\partial_t(\widehat{v}_h - v_h), \widehat{v}_h - v_h) = -(\partial_t(v-\widehat{v}_h), \widehat{v}_h - v_h)$$

Next set

$$E_h(t) = \frac{1}{2}\left(a(\widehat{p}_h - p_h, \widehat{p}_h - p_h) + (\widehat{v}_h - v_h, \widehat{v}_h - v_h)\right)(t)$$

Since for some constant $C > 0$, we have

$$E_h^{1/2}(t) \geq C\left(\|\widehat{p}_h - p_h\|_H^2(t) + \|\widehat{v}_h - v_h\|_M^2(t)\right)^{1/2}$$

It follows that

$$(63)\qquad \frac{dE_h^{1/2}}{dt}(t) \leq C'\|\partial_t(v-\widehat{v}_h)\|_M(t) \equiv C'\|\partial_t\delta_h\|_M(t)$$

where we set $\delta_h = v - \widehat{v}_h$ and from the choice (54) for the initial conditions

$$E_h(0) = 0$$

It is then easy to see that (63) leads to (57) (from the orthogonality $X_h^s \perp X_h^r$, $\|q_h\|_H^2 = \|q_h^s\|_H^2 + \|q_h^r\|_H^2$, $\forall q_h \in X_h$) .

• Estimation (58) : to get the estimate on the X norm, we recall that forall $\eta_h^s \in X_h^s$ we have $\eta_h^s = \eta_1 + \eta_2$ with $\eta_1 \in KerB_h$ and $\eta_2 \in (KerB_h)^\perp$ so

$$\|\eta_h^s\|_X^2 = \|\eta_1\|_H^2 + \|\eta_2\|_X^2$$

(recalling that $\|\eta_1\|_X = \|\eta_1\|_H \ \ \forall \eta_1 \in KerB_h$). We set $\eta_h = \widehat{p}_h - p_h$, the term $\|\eta_1\|_H$ is already estimated from the inequality (57), therefore in order to get the second inequality (58) we only need to estimate $\|\eta_2\|_X$. To do so, we start from reminding that the inf-sup condition (hypothesis (H2)) is equivalent to :

$$(64)\qquad \left\{\begin{array}{l} \text{there exists a constant } C > 0, \text{ independent of } h, \text{ such that} \\ \forall\ q_h^s \in X_h^s, \quad \displaystyle\sup_{w_h \in M_h} \frac{b(w_h, q_h^s)}{\|w_h\|_M} \geq C\ \|q_h^s\|_{X/KerB_h} \end{array}\right.$$

we also know that $\|\eta_2\|_X = \|\eta_h^s\|_{X/KerB_h}$. Thus by taking $q_h^s = \eta_2$ in (64) we get

$$\sup_{w_h \in M_h} \frac{b(w_h, \eta_2)}{\|w_h\|_M} \geq C\ \|\eta_2\|_X$$

using now the second equation of (61) we obtain

$$\|\eta_2\|_X \leq C' \{\|\partial_t(v - \widehat{v}_h)\|_M + \|\partial_t(\widehat{v}_h - v_h)\|_M\} \tag{65}$$

Till now, we only have used the C^1 regularity of the solution. In order to bound $\|\partial_t(\widehat{v}_h - v_h)\|_M$, we need C^2. Indeed, we want to apply (57) replacing v_h by $\partial_t v_h$, $\widehat{v}_h$ by $\partial_t \widehat{v}_h$ and so on... More precisely, we have

$$\|\partial_t(\widehat{v}_h - v_h)\|_M(t) \leq C \int_0^t \|\partial_t^2 \delta_h\|_M(s)\, ds \tag{66}$$

Finally, combining (65), (66) we get

$$\|\eta_2\|_X(t) \leq C' \left\{ \|\partial_t \delta_h\|_M(t) + \int_0^t \|\partial_t^2 \delta_h\|_M(s)\, ds \right\}$$

and the proof is achieved.□

Proof of Theorem 8. We combine results given in lemma 4 and in lemma 5.

- Estimates in norms H and M :

  We have

$$\|p - p_h^s\|_H + \|p_h^r\|_H \leq (\|p - \widehat{p}_h^s\|_H + \|\widehat{p}_h^r\|_H) + (\|\widehat{p}_h^s - p_h^s\|_H + \|\widehat{p}_h^r - p_h^r\|_H) \tag{67}$$

The first term in the right hand side is bounded by $|||p - \widehat{p}_h|||$ and thus by $\| [(p,u) - \Pi_h(p,u)] \|$. Assuming that $(p,v) \in \mathcal{C}^{1,0} \times C^1(0,T;M)$, it can be estimated by (cf (51))

$$|||p - \widehat{p}_h|||\,(t) \leq C\ \mathcal{N}_h(p,u)(t) \tag{68}$$

The second term is estimated using (57) of lemma 5. This requires to estimate $\|\partial_t \delta_h\|_M$. For this, we use (52), for $k = 2$, which requires $(p,v) \in C^2(0,T;X) \times C^1(0,T;M)$. We get

$$\|\partial_t(v - \widehat{v}_h)\|_M(s) \leq C\ \mathcal{N}_h\left(\partial_t^2 p, \partial_t v\right)(s) \tag{69}$$

(67), (68) and (69) lead to the first inequality of (55). Now, for v, we write :

$$\|v - v_h\|_M \leq \|v - \widehat{v}_h\|_M + \|\widehat{v}_h - v_h\|_M$$

We apply (52), for $k = 1$ and get

$$\|v - \widehat{v}_h\|_M \leq C\ \mathcal{N}_h\,(\partial_t p, v)$$

Using again estimate (57) for bounding $\|\widehat{v}_h - v_h\|_M$, we easily get the second inequality of (55).

- Estimate in norm X :

$$\|p - p_h^s\|_X \leq \|p - \widehat{p}_h^s\|_X + \|\widehat{p}_h^s - p_h^s\|_X$$

The first term is again bounded by $|||p - \widehat{p}_h|||$ and can be estimated from (51) of lemma 4. For the second term, we now use (58) of lemma 5 which requires $(p,v) \in \mathcal{C}^{2,1} \times C^2(0,T;M)$ and $(p_h, v_h) \in C^2(0,T;X_h) \times C^2(0,T;M_h)$. In the right hand side of (58) appears the second derivative of $v - \widehat{v}_h$. We thus use estimate (52), for $k = 3$, which requires $(p,v) \in C^3(0,T;X) \times C^2(0,T;M)$ and we get (56). □

4.2 Application to the approximation of the anisotropic wave equation with the new finite element, $Q_{k+1}^{div} - Q_k$

We come back to the original problem described in section 2. We consider the approximate spaces given in (17). From section 3.3 we know that the $Q_{k+1}^{div} - Q_k$ element enters the abstract framework. It is then straightforward to apply theorem 8 and we get :

Theorem 9 *Let (p,v) be the solution of (5), (4) and $u \in C^2(0,T;M)$ the primitive of v, satisfying at intial time*

$$\frac{du}{dt} = v \ ; \ \ B^t u(0) = -\mathcal{A}p_0$$

Let (p_h, u_h) be the solution of the approximate problem (9) with initial conditions :

$$(p_{0,h}, v_{0,h}) = \Pi_h(p_0 \ , \ v_0)$$

- Convergence in L^2 norm :

(i) If $(p,v) \in C^2(0,T;X) \times C^1(0,T;M)$, forall $t \in [0,T]$

$$\|p - p_h^s\|_H(t) \to 0 \ ; \ \ \|p_h^r\|_H(t) \to 0 \ ; \ \ \|v - v_h\|_M(t) \to 0$$

(ii) Further, if we assume that the solution $(p,u) \in C^2(0,T;H^{k+1}(div,\Omega)) \times C^2(0,T;H^{k+1}(\Omega))$,

$$\|p - p_h^s\|_H(t) + \|p_h^r\|_H(t) + \|v - v_h\|_M(t) \leq C_1(t)h^k$$

with $C_1(t) = O(\|p\|_{C^2(0,t;H^{k+1}(div,\Omega))} + \|\mathcal{A}p\|_{C^2(0,t;H^{k+1}(\Omega))} + \|u\|_{C^2(0,t;H^{k+1}(\Omega))})$.

- Convergence in $H(div)$ norm :

(iii) If $(p,v) \in C^3(0,T;X) \times C^2(0,T;M)$, forall $t \in [0,T]$

$$\|p - p_h^s\|_X(t) \to 0$$

(iv) Further, if we assume that the solution $(p,u) \in C^3(0,T;H^{k+1}(div,\Omega)) \times C^3(0,T;H^{k+1}(\Omega))$,

$$\|p - p_h^s\|_X(t) \leq C_2(t)h^k$$

with $C_2(t) = O(\|p\|_{C^3(0,t;H^{k+1}(div,\Omega))} + \|\mathcal{A}p\|_{C^3(0,t;H^{k+1}(\Omega))} + \|u\|_{C^3(0,t;H^{k+1}(\Omega))})$

Proof of theorem 9 : We apply theorem 8 which relates the errors to quantities as $\mathcal{N}_h(\partial_t^m p, \partial_t^l u)$, i.e., to the error due to the approximation of space $H(div,\Omega)$ with the Raviart-Thomas $RT_{[k]}$ space in norm L^2 and in norm $H(div)$ and due to the approximation of space $L^2(\Omega)$ with the Q^k discontinuous elements. □

5 Extension in 3D

We will present in this section the new finite element in 3D, only in the lowest order, the extension to higher orders being similar to the 2D case.

Let Ω be a bounded domain of $\mathbb{R}^3$ and consider a regular mesh $(\mathcal{T}_h)$ of Ω with cubic elements (K) of edge $h > 0$. We introduce the approximation spaces $X_h \subset X (= H(div,\Omega))$ and $M_h \subset M (= L^2(\Omega))$ defined by (8) with :

$$\hat{X} = Q_1 \times Q_1 \times Q_1$$

$$\hat{M} = Q_0$$

Our aim now is to obtain error estimates for the approximation problem (23) by applying the abstract Theorem 1 as in the 2D case. Defining first X_h^s as the lowest order $(RT_{[0]})$ element :

$$X_h^s = \{q_h \in X \ / \ \forall K \in \mathcal{T}_h, \ q_h|_K \in P_{1,0,0}(K) \times P_{0,1,0}(K) \times P_{0,0,1}(K)\}$$

introduced by Nedelec in [15], we see that we can use the well known properties of $X_h^s - M_h$ (cf. [7], [15]), in order to obtain the same error estimates as in Theorem 2. The only point which is not trivial is the orthogonal

decomposition of X_h.
In order to describe the space $X_h^r = (X_h^s)^\perp$ we denote as illustrated in figure (13),

$$(\vec{\phi}_{i,j+\frac{1}{2},k+\frac{1}{2}},\ \vec{\phi}_{i+\frac{1}{2},j,k+\frac{1}{2}},\ \vec{\phi}_{i+\frac{1}{2},j+\frac{1}{2},k})$$

the base functions in $RT_{[0]}$.

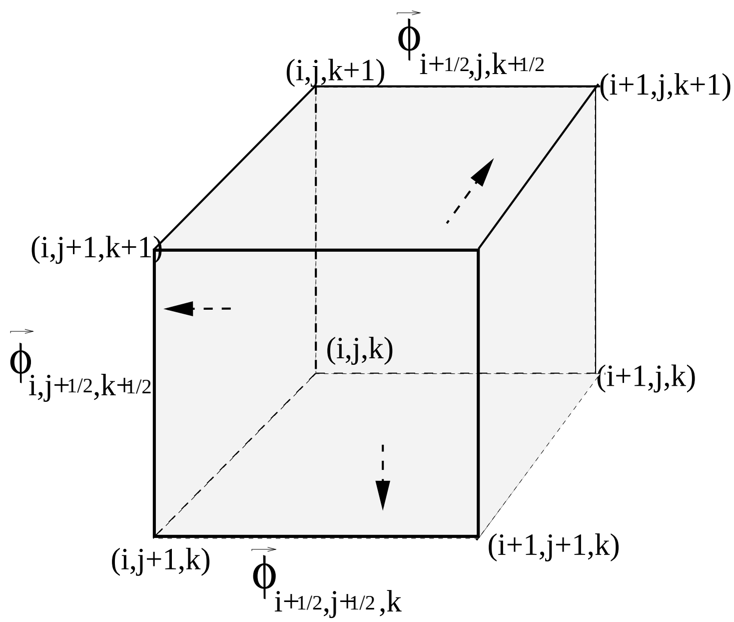

Figure 13: The base functions in $RT_{[0]}$.

Which can be written in the following form

$$\vec{\phi}_{i,j+\frac{1}{2},k+\frac{1}{2}} = \begin{pmatrix} \phi_{i,j+\frac{1}{2},k+\frac{1}{2}} \\ 0 \\ 0 \end{pmatrix},\ \vec{\phi}_{i+\frac{1}{2},j,k+\frac{1}{2}} = \begin{pmatrix} 0 \\ \phi_{i+\frac{1}{2},j,k+\frac{1}{2}} \\ 0 \end{pmatrix},\ \vec{\phi}_{i+\frac{1}{2},j+\frac{1}{2},k} = \begin{pmatrix} 0 \\ 0 \\ \phi_{i+\frac{1}{2},j+\frac{1}{2},k} \end{pmatrix}$$

It is then easy to prove that the space X_h^r can be generated by the following functions :

$$\vec{\psi}_{i,j+\frac{1}{2},k+\frac{1}{2}} = \begin{pmatrix} \phi_{i,j+\frac{1}{2},k+\frac{1}{2}}(y-y_{i+\frac{1}{2}})(z-z_{k+\frac{1}{2}}) \\ 0 \\ 0 \end{pmatrix},$$

$$\vec{\psi}_{i+\frac{1}{2},j,k+\frac{1}{2}} = \begin{pmatrix} 0 \\ \phi_{i+\frac{1}{2},j,k+\frac{1}{2}}(x-x_{i+\frac{1}{2}})(z-z_{k+\frac{1}{2}}) \\ 0 \end{pmatrix},$$

$$\vec{\psi}_{i+\frac{1}{2},j+\frac{1}{2},k} = \begin{pmatrix} 0 \\ 0 \\ \phi_{i+\frac{1}{2},j+\frac{1}{2},k}(x-x_{i+\frac{1}{2}})(y-y_{j+\frac{1}{2}}) \end{pmatrix}$$

Moreover we have $X_h^r \subset V_h$.

Conclusion

We have presented in this paper a family of mixed finite elements leading to an explicit time discretisation scheme for the anisotropic wave equation. The generalization of these elements to the linear elastodynamic problem involves a main difficulty concerning the symmetry of the stress tensor. An answer to this problem will be given in a next paper, where inspired from the work that we have presented here, we will describe the construction of a new family of mixed finite elements for linear elasticity, leading also to an explicit time discretisation scheme.

References

[1] D. N. Arnold, F. Brezzi, and J. Douglas. A new mixed finite element for plane elasticity. *Japan J. Appl. Math.*, 1:347–367, 1984.

[2] G.A. Baker and V.A. Dougalis. The effect of quadrature errors on finite element approximations for the second-order hyperbolic equations. *SIAM jour. numer. anal.*, 13(4):577–598, 1976.

[3] E. Bécache, P. Joly, and C. Tsogka. Eléments finis mixtes et condensation de masse en élastodynamique linéaire. (i) construction. *C.R. Acad. Sci. Paris*, t. 325, Série I:545–550, 1997.

[4] E. Bécache, P. Joly, and C. Tsogka. Etude d'un nouvel élément fini mixte permettant la condensation de masse. *C.R. Acad. Sci. Paris*, t. 324, Série I:1281–1286, 1997.

[5] J.P. Bérenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Comp. Physics.*, **114**:185–200, 1994.

[6] F. Brezzi. On the existence uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *Revue Francaise d'Automatique, Informatique et Recherche Opérationnelle*, R-2:129–151, 1974.

[7] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods.* Springen–Verlag, 1991.

[8] F. Collino, P. Joly, and F. Millot. Fictitious domain method for unsteady problems: Application to electromagnetic scattering. Technical Report 2963, INRIA, Août 1996.

[9] P.J. Davis and P. Rabinowitz. *Methods of Numerical Integration.* Academic Press, Orlando, 1985.

[10] T. Dupont. l^2-estimate for galerkin methods for second order hyperbolic equations. *SIAM jour. numer. anal.*, 10(5), 1973.

[11] A. Elmkies and P. Joly. Éléments finis et condensation de masse pour les équations de maxwell: le cas 2d. Technical Report 3035, INRIA, Novembre 1996.

[12] N. Tordjman G. Cohen, P. Joly. Construction and analysis of higher order finite elements with mass lumping for the wave equation. In SIAM, editor, *Proccedings of the second Int. Conf. on Mathematical and Numerical Aspects of Wave Propagation*, pages 152–160, U. of Dalaware, USA, 1993.

[13] P. Monk G. Cohen. Gauss point mass lumping schemes for maxwell's equations.

[14] Mary E. Morley. A family of mixed finite elements for linear elasticity. *Numer. Math.*, 55:633–666, 1989.

[15] J.C. Nédélec. Mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 35:315–341, 1980.

[16] J.C. Nédélec. A new family of mixed finite elements in $\mathbb{R}^3$. *Numer. Math.*, 50:57–81, 1986.

[17] P.A. Raviart and J.M. Thomas. A mixed finite element method for 2nd order elliptic problems. Number 606, pages 292–315. Proc. of math. aspects on the finite element method, 1977.

[18] J.E. Roberts and J.M. Thomas. *Mixed and Hybrid Methods, in Handbook of Numerical Analysis.* 1989.

[19] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53:513–538, 1988.

[20] N. Tordjman. *Eléments finis d'ordre élevé avec condensation de masse pour l'équation des ondes.* PhD thesis, Univ. Paris IX, 1995.

[21] V.Girault and R. Glowinski. Error analysis of a fictitious domain method applied to a dirichlet problem. *Japan J. Indust. Appl. Math.*, 12(3):487–514, 1995.

Unit´e de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399